Q1: Why does the DNA sequence have more dots than the protein sequence plot?

More amino acids than nucleotides. AGCT to AGCT is somple so it’s just mathematically easier to get a hit.

Q2: How can we increase the signal to noise ratio?

Window size?

Q3: What does a 'Match stringency' larger than 'Window size' yield and why?

No results

Q4: What would off-diagonal runs of dots represent?

Noise

Q5: What are the major weaknesses of this approach?

The longer, the harder to look at.

Q9:

No

Q10:

First all are flies, almost all Drosophila species

Q11:

desert hedgehog protein preproprotein [Homo sapiens]

Q12:

The hedgehog gene family encodes signaling molecules that play an important role in regulating morphogenesis.

Q13:

Stomoxys calcitrans

Q14:

Only Drosophila

Q15:

Yes

Q16:

Just one

Q17: